


# Abdul Mannan

## Data Scientist

### CONTACT

www.abdulmannan133@gmail.com 

+923247562530 

RYK, Pakistan 

[LinkedIn](#) 

### EDUCATION

B.S.

Software  
Engineering

Islamia University of  
Bahawalpur, Pakistan

September 2020 - April 2024

### SKILLS

Programming: Python (Scikit-learn,  
Pandas), SQL, JS, C++

Data Visualization: PowerBi, MS Excel,  
Plotly

Modeling: Regression, Classification  
, decision tree, NLP

Databases: Postgres, MongoDB,  
MySQL

### SOFT SKILLS

Communication  
Problem-Solving  
Time Management  
Teamwork  
Adaptability  
Attention to Detail

### HOBBIES

Cycling  
Swimming  
Travelling

### Projects

#### Houses price prediction Model

- Developed a Houses Price Prediction Model using XGBoost
- Conducted comprehensive data preprocessing, feature engineering, and exploratory data analysis to enhance model performance and interpretability.
- Created data visualizations and interpreted feature importance to explain the model's predictions to stakeholders.

#### Restaurant Reviews Classification Model

- Developed a Restaurant Reviews Classification Model using NLP techniques and Gaussian Naive Bayes (GaussianNB) algorithm, achieving an 73% accuracy.
- Preprocessed and cleaned text data, including tasks such as tokenization, stop word removal, and stemming/lemmatization, to improve the quality of input features.
- Engineered textual features like TF-IDF and word embeddings to capture semantic meaning and sentiment in reviews.
- Conducted exploratory data analysis (EDA) and visualizations to gain insights into the dataset's characteristics, which informed feature engineering decisions.

#### Social Network Ads Classifier

- Developed a Social Network Ads Classifier using the K-Nearest Neighbors (K-NN) algorithm, achieving an 93% accuracy
- Conducted data preprocessing and feature scaling to ensure accurate and efficient classification of social media users' ad interactions.
- Engineered relevant features and performed exploratory data analysis (EDA) to better understand the data's characteristics and improve model performance.
- Implemented cross-validation techniques to assess the model's generalization and robustness, ensuring reliable predictions for new data.

### Certifications

Data Manipulation in Python: Master Python, Numpy & Pandas

Python Data Science Toolbox

IBM Full Stack Software Developer

### Language

English (proficient)

Urdu (proficient)

Punjabi (Native)

